

以下译文仅供参考。如果本译文与最后更新的英文版本之间存在差异、不一致或冲突（包括由于翻译延误），则以英文版本为准。

Amazon SageMaker 服务水平协议

最后更新日期：2022 年 5 月 4 日

本 Amazon SageMaker 服务水平协议（“SLA”）是一项在 AWS 客户协议或您与我们签订的其他适用于您使用我们服务的协议（“协议”）的条款下管理 Amazon SageMaker 使用的政策。本 SLA 单独适用于每个使用 Amazon SageMaker 的账户。除非另有规定，否则本 SLA 受协议条款的约束，且专门术语应具有协议中规定的含义。我们保留根据协议更改本 SLA 条款的权利。

本 SLA 适用于 Amazon SageMaker 的以下功能：

- “在线推理” - Amazon SageMaker 的调用端点 (InvokeEndpoint) API，用于由多个实例支持的模型端点。
- “批量转换” - Amazon SageMaker 的创建转换任务 (CreateTransformJob) API 和停止转换任务 (StopTransformJob) API。

服务承诺

AWS 将尽商业上合理的努力，确保在任何月度账单周期内，每个 AWS 区域的 Amazon SageMaker 达到每月正常运行时间百分比（定义见下文）（“服务承诺”）。如果 Amazon SageMaker 未达到服务承诺，您将有资格获得如下所述的服务抵扣额度。

定义

- “服务抵扣额度”是一种美元抵扣额度，以下文所述的方式计算，我们可以将该额度退回符合条件的账户。
- 给定 AWS 区域的“每月正常运行时间百分比”是指月度账单周期中所有 5 分钟时间段的可用性平均值。每月正常运行时间百分比不包括任何 Amazon SageMaker SLA 例外情况（定义见下文）直接或间接导致的停机时间。
- “可用性”是指按每 5 分钟时间段计算的，Amazon SageMaker 处理的未因错误（定义见下文）而失败且仅与所提供的 Amazon SageMaker 功能相关的请求（定义见下文）所占的百分比。如果您在给定的 5 分钟内未发出任何请求，则假定该时间段内 100% 可用。
- “请求”是指：
 - 对于“在线推理”而言，是指通过直接调用 InvokeEndpoint API 来调用 Amazon SageMaker 功能。
 - 对于“批量转换”而言，是指通过直接调用 CreateTransformJob 和 StopTransformJob API 来调用 Amazon SageMaker 功能。

以下译文仅供参考。如果本译文与最后更新的英文版本之间存在差异、不一致或冲突（包括由于翻译延误），则以英文版本为准。

- “错误”是指返回 500 或 503 错误代码的任何请求。

服务抵扣额度

服务抵扣额度按照在每月正常运行时间百分比处于下表所示范围的月度账单周期内，您为给定 AWS 区域的 Amazon SageMaker “在线推理”或“批量转换”（按适用情形）支付的以下费用的百分比计算：

对于“在线推理”：

每月正常运行时间百分比	服务抵扣额度百分比
低于 99.95% 但高于或等于 99%	10%
低于 99%	25%
低于 95%	100%

对于“批量转换”：

每月正常运行时间百分比	服务抵扣额度百分比
低于 99.9% 但高于或等于 99%	10%
低于 99%	25%
低于 95%	100%

任何服务抵扣额度只适用于抵扣您在未来须支付的 Amazon SageMaker “在线推理”或“批量转换”（按适用情形）款项。我们可能会自行决定向您的信用卡（即在发生不可用状况的账单周期内，您用来付款的信用卡）发放服务抵扣额度。服务抵扣额度将不会使您有权从 AWS 获得任何退款或其他付款。服务抵扣额度仅在相应月度账单周期的抵扣金额超过一美元 (\$1 USD) 时适用并发放。服务抵扣额度不得转让或应用于任何其他账户。除非协议另有规定，否则对于任何 Amazon SageMaker 不可用、性能不佳的情况，或我们未能提供 Amazon SageMaker 的其他情况，您享有的唯一且排他的救济限于根据本 SLA 的条款获得服务抵扣额度（如符合资格要求）。

以下译文仅供参考。如果本译文与最后更新的英文版本之间存在差异、不一致或冲突（包括由于翻译延误），则以英文版本为准。

抵扣额度申请和支付程序

要获得服务抵扣额度，您需要在 **AWS 支持中心** 创建问题单，提交申请。为了符合条件，我们必须在事件发生后的第二个账单周期结束前收到抵扣额度申请，且该申请必须包括：

- (i) 主题行中的“**SLA 抵扣额度申请**”字样；
- (ii) 您申请服务抵扣额度的账单周期和 **AWS 区域**，以及该账单周期内该 **AWS 区域** 的每月正常运行时间百分比，以及在整個账单周期内该 **AWS 区域** 可用性低于 **100%** 的每个 **5 分钟** 时间段的具体日期、次数和可用性；
- (iii) 记录您所述故障之错误的申请日志（请删除日志中的任何机密或敏感信息，或用星号替代）。

如果我们确认了此类抵扣额度申请的每月正常运行时间百分比，且其低于服务承诺，则我们将在抵扣额度申请发生月份后的一个账单周期内向您发放服务抵扣额度。如果您未能按照上述要求提供抵扣额度申请和其他信息，您将失去获得服务抵扣额度的资格。

Amazon SageMaker SLA 例外情况

服务承诺不适用于因以下原因导致的任何 **Amazon SageMaker** 不可用、暂停或终止，或任何其他 **Amazon SageMaker** 性能问题：

- (i) 因超出我们合理控制范围的因素（包括任何不可抗力事件、互联网接入或超出 **Amazon SageMaker** 分界点的相关问题）而导致；
- (ii) 因您的任何自愿作为或不作为（例如，扩展配置容量、错误配置安全组、**VPC** 配置或凭据设置、禁用加密密钥或使加密密钥无法访问等）而导致；
- (iii) 因您未遵循 **AWS** 网站上 **Amazon SageMaker** 文档中所述的最佳实践而导致；
- (iv) 因您的设备、软件或其他技术而导致；或
- (v) 因我们根据协议暂停和终止您使用 **Amazon SageMaker** 的权利而导致（统称为“**Amazon SageMaker SLA 例外情况**”）。

如果可用性受到我们在计算每月正常运行时间百分比时明确使用的因素以外的因素影响，则我们在发放服务抵扣额度时可能会酌情考虑这些因素。